OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/998,904

DATE: 06/27/2002
TIME: 19:18:55

Input Set : N:\Crf3\06112002\I998904.raw
Output Set: N:\CRF3\06272002\I998904.raw

```
 1 <110> APPLICANT: The University of Texas Board of Regents
 2 <120> TITLE OF INVENTION: Prediction of Disease-Causing Alleles from Sequence
 3       Context
 4 <130> FILE REFERENCE: 119929-1037
 5 <140> CURRENT APPLICATION NUMBER: US/09/998,904
 6 <141> CURRENT FILING DATE: 2002-06-04
 7 <160> NUMBER OF SEQ ID NOS: 12
 8 <170> SOFTWARE: PatentIn version 3.1
10 <210> SEQ ID NO: 1
11 <211> LENGTH: 101
12 <212> TYPE: DNA
13 <213> ORGANISM: mammal
14 <220> FEATURE:
15 <221> NAME/KEY: misc_feature
16 <222> LOCATION: (51)..(51)
17 <223> OTHER INFORMATION: n = a or g
18 <400> SEQUENCE: 1
```

W--> 19   gctgggccaa gctctacagc ttggtgatct gggggtgtac gctgctcctg ngctcaccca   60
20   tgctggtgtt ccggaccatg aaggagtaca gcgatgaggg c   101

```
22 <210> SEQ ID NO: 2
23 <211> LENGTH: 101
24 <212> TYPE: DNA
25 <213> ORGANISM: mammal
26 <220> FEATURE:
27 <221> NAME/KEY: misc_feature
28 <222> LOCATION: (51)..(51)
29 <223> OTHER INFORMATION: n = c or g
30 <400> SEQUENCE: 2
```

W--> 31   gctggtgttc cggaccatga aggagtacag cgatgagggc cacaacgtca ncgcttgtgt   60
32   catcagctac ccatccctca tctgggaagt gttcaccaac a   101

```
34 <210> SEQ ID NO: 3
35 <211> LENGTH: 101
36 <212> TYPE: DNA
37 <213> ORGANISM: mammal
38 <220> FEATURE:
39 <221> NAME/KEY: misc_feature
40 <222> LOCATION: (51)..(51)
41 <223> OTHER INFORMATION: n = a or g
42 <400> SEQUENCE: 3
```

W--> 43   gggccaagct ctacagcttg gtgatctggg ggtgtacgct gctcctgagc ncacccatgc   60
44   tggtgttccg gaccatgaag gagtacagcg atgagggcca c   101

```
46 <210> SEQ ID NO: 4
47 <211> LENGTH: 101
```

```
       48 <212> TYPE: DNA
       49 <213> ORGANISM: mammal
       50 <220> FEATURE:
       51 <221> NAME/KEY: misc_feature
       52 <222> LOCATION: (51)..(51)
       53 <223> OTHER INFORMATION: n = a or g
       54 <400> SEQUENCE: 4
W--> 55        ctgcccttct gggccatcac catctccaac aacttcgact ggctctttgg ngagacgctc      60
       56        tgccgcgtgg tgaatgccat tatctccatg aacctgtaca g                       101
       58 <210> SEQ ID NO: 5
       59 <211> LENGTH: 101
       60 <212> TYPE: DNA
       61 <213> ORGANISM: mammal
       62 <220> FEATURE:
       63 <221> NAME/KEY: misc_feature
       64 <222> LOCATION: (51)..(51)
       65 <223> OTHER INFORMATION: n = c or t
       66 <400> SEQUENCE: 5
W--> 67        cttctgggcc atcaccatct ccaacaactt cgactggctc tttggggaga ngctctgccg      60
       68        cgtggtgaat gccattatct ccatgaacct gtacagcagc a                       101
       70 <210> SEQ ID NO: 6
       71 <211> LENGTH: 101
       72 <212> TYPE: DNA
       73 <213> ORGANISM: mammal
       74 <220> FEATURE:
       75 <221> NAME/KEY: misc_feature
       76 <222> LOCATION: (51)..(51)
       77 <223> OTHER INFORMATION: n = g or a
       78 <400> SEQUENCE: 6
W--> 79        ggacaccggc caccctccgc gccacccacc ctcgctttct ccggcttcct ntggcccagg      60
       80        cgccgcgcgg acccggcagc tgtctgcgca cgccgagctc c                       101
       82 <210> SEQ ID NO: 7
       83 <211> LENGTH: 101
       84 <212> TYPE: DNA
       85 <213> ORGANISM: mammal
       86 <220> FEATURE:
       87 <221> NAME/KEY: misc_feature
       88 <222> LOCATION: (51)..(51)
       89 <223> OTHER INFORMATION: n = g or a
       90 <400> SEQUENCE: 7
W--> 91        ctgtctcagc agtggacagc gggcatgggt ctgctgatgg cgctcatcgt nctgctcatc      60
       92        gtggcgggca atgtgctggt gatcgtggcc atcgccaaga c                       101
       94 <210> SEQ ID NO: 8
       95 <211> LENGTH: 101
       96 <212> TYPE: DNA
       97 <213> ORGANISM: mammal
       98 <220> FEATURE:
       99 <221> NAME/KEY: misc_feature
      100 <222> LOCATION: (51)..(51)
```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/998,904

DATE: 06/27/2002
TIME: 19:18:55

Input Set : N:\Crf3\06112002\I998904.raw
Output Set: N:\CRF3\06272002\I998904.raw

```
      101 <223> OTHER INFORMATION: n = t or c
      102 <400> SEQUENCE: 8
W--> 103       cgggcatggg tctgctgatg gcgctcatcg tgctgctcat cgtggcgggc natgtgctgg       60
      104       tgatcgtggc catcgccaag acgccgcggc tgcagacgct c                          101
      106 <210> SEQ ID NO: 9
      107 <211> LENGTH: 101
      108 <212> TYPE: DNA
      109 <213> ORGANISM: mammal
      110 <220> FEATURE:
      111 <221> NAME/KEY: misc_feature
      112 <222> LOCATION: (51)..(51)
      113 <223> OTHER INFORMATION: n = a or g
      114 <400> SEQUENCE: 9
W--> 115       ctgctgatgg cgctcatcgt gctgctcatc gtggcgggca atgtgctggt natcgtggcc       60
      116       atcgccaaga cgccgcggct gcagacgctc accaacctct t                         101
      118 <210> SEQ ID NO: 10
      119 <211> LENGTH: 101
      120 <212> TYPE: DNA
      121 <213> ORGANISM: mammal
      122 <220> FEATURE:
      123 <221> NAME/KEY: misc_feature
      124 <222> LOCATION: (51)..(51)
      125 <223> OTHER INFORMATION: n = t or c
      126 <400> SEQUENCE: 10
W--> 127       tcgccaagac gccgcggctg cagacgctca ccaacctctt catcatgtcc ntggccagcg       60
      128       ccgacctggt catggggctg ctggtggtgc cgttcggggc c                         101
      130 <210> SEQ ID NO: 11
      131 <211> LENGTH: 101
      132 <212> TYPE: DNA
      133 <213> ORGANISM: mammal
      134 <220> FEATURE:
      135 <221> NAME/KEY: misc_feature
      136 <222> LOCATION: (51)..(51)
      137 <223> OTHER INFORMATION: n = a or g
      138 <400> SEQUENCE: 11
W--> 139       cgtggtgtgg ggccgctggg agtacggctc cttcttctgc gagctgtgga nctcagtgga       60
      140       cgtgctgtgc gtgacggcca gcatcgagac cctgtgtgtc a                         101
      142 <210> SEQ ID NO: 12
      143 <211> LENGTH: 100
      144 <212> TYPE: DNA
      145 <213> ORGANISM: mammal
      146 <220> FEATURE:
      147 <221> NAME/KEY: misc_feature
      148 <222> LOCATION: (51)..(51)
      149 <223> OTHER INFORMATION: n = t or c
      150 <400> SEQUENCE: 12
      151       ttccgctacc agagcctgct gacgcgcgcg cgggcgcggg gcctcgtgtg accgtgtggg       60
      152       ccatctcggc cctggtgtcc ttcctgccca tcctcatgca                           100
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 06/27/2002
PATENT APPLICATION: US/09/998,904   TIME: 19:18:56

Input Set : N:\Crf3\06112002\I998904.raw
Output Set: N:\CRF3\06272002\I998904.raw

<u>Please Note:</u>
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 51
Seq#:2; N Pos. 51
Seq#:3; N Pos. 51
Seq#:4; N Pos. 51
Seq#:5; N Pos. 51
Seq#:6; N Pos. 51
Seq#:7; N Pos. 51
Seq#:8; N Pos. 51
Seq#:9; N Pos. 51
Seq#:10; N Pos. 51
Seq#:11; N Pos. 51